Copper Price Discovery on Comex, the LME and the SHFE, 2001-2013

Isabel Figuerola-Ferretti
Christopher L. Gilbert
Jieqin Yan

Initial draft: 2 November 2013
This revision: 23 February 2014

Abstract

Over the past two decades, China has come to dominate international commerce in copper. The importance of the Shanghai Futures Exchange (SHFE) has increased in response to this development. We look at the distribution of price discovery between the SHFE and the two historically important copper futures exchanges, Comex and the LME. The results indicate that it is Comex, followed by the SHFE, not the LME which plays the most important role in copper price discovery. We also highlight a number of problems associated with both the calculation and interpretation of the standard IS and PT price discovery measures when used to look at overlapping price change on non-synchronous markets. The results offer a clearer interpretation in terms of trading slots (European, North American and Asian trading days) than in terms of exchanges.

The initial draft of this paper was prepared for Second International Conference on Futures and other Derivative Markets (ICFDM), Renmin University, Beijing, 8-9 November 2013. The authors are responsible for the entire content. We are grateful to the ICFDM discussant, Andrew Caminschi, for comments on the conference paper and to Wan Ling and Li Chunlan of the CRU Group for discussion of the Chinese copper market.

Addresses for correspondence: Figeurola-Ferretti (ifgarrig@emp.uc3m.es) Departamento de Economía de la Empressa, Universidad Carlos III de Madrid, Calle Madrid 126, 28093 Getafe (Madrid); Gilbert (christopher.gilbert@unitn.it) Department of Economics and Management, Università degli Studi di Trento, via Inama 5, 38122 Trento, Italy (corresponding author); Yan (jyan@planetfinance.org) PlaNet Finance Beijing Office, Bldg 5 MOMA, No.1 Xiangheyuan Rd, Beijing, P.R.China.

1. Introduction

This paper analyses equilibrium price dynamics in the copper market. Copper futures are traded at non-synchronous times in three different continents: Europe, North America and Asia. Copper exchange trading started in London as one of original London Metal Exchange (LME) metals and has been traded continuously since the reopening of the exchange after the Second World War. The Commodity Exchange of New York (Comex),[1] which has traded high grade copper since 1988, has traditionally been the principal competitor for the LME. With the advent of sustained growth in China, the Shanghai Futures Exchange (SHFE) has become a significant player in the copper market to the extent that volumes traded in SHFE has now reached comparable levels to those registered by its competitors. We consider the relative contributions of the three markets to price discovery over the twelve year period 2001-13.

Price discovery is one of the most important functions of futures markets. Exchange quoted prices are widely used by commodity index investors as well as firms engaged in the production and consumption of commodities. When several prices are quoted, it is important for all concerned to understand which of these most efficiently reflects the underlying market fundamentals. A contract fails to contribute in a substantial way to price discovery will be a follower rather than a leader and this will undermine its long term viability and prospects for survival. Regulators need to understand whether the markets for which they are concerned are competing effectively and how the different markets interact with each other. They will wish to ensure that regulation enhances rather than impedes discovery in the exchanges they regulate.

The periods of active trading in the three copper futures markets only partially overlap. Comex closing prices are determined at 13:00 in local time. The SHFE close is 15:00 local time, equivalent to 02:00 in New York. Official LME copper prices are determined at 12:30 local time, equivalent to 07:30 in New York. Unofficial prices are determined at 16:15 local time equivalent to 10:15 in New York. Trading activity therefore tends to move round the world depending on which market is active. There is no single period in which all three markets are actively trading. Although it is always possible to trade on the Comex and LME electronic markets at any time, these platforms generally exhibit relatively low liquidity outside North American and European working hours.

[1] Now part of the CME Group.

Previous studies on market integration in distinct geographical generally rely on high frequency equity market data for overlapping trading hours (Hupperets and Menkeveld, 2002; Pascual et al., 2001). Since the three markets we consider are only partially overlapping, we analyze official and closing prices. The price discovery literature commonly applies the standard Gonzalo-Granger (1995) Permanent-Transitory (PT) and Hasbrouck (1995) Information Shares (IS) procedures, both of which rely on an estimated Vector AutoRegression (VAR) model. These approaches have previously been applied to the Comex, LME and SHFE copper futures markets by Hua et al. (2010) who find that over 1998-2008, the LME market remained dominant while the SHFE market grew in importance.

The IS and PT procedures generate fairly similar results in the standard context of two simultaneously traded markets. Building on the contribution of Lieberman et al. (1999), we show that the IS and PT approaches have radically different interpretations and implications when applied to non-synchronous markets. Our analysis clarifies the differences between the IS and PT discovery measures. Depending on the econometric specification, discovery measures can relate to markets, time slots or latent factors. Furthermore, application of these procedures becomes more complicated once moves beyond the standard case of two markets.

The principal substantive conclusion from this study is the important role in price discovery played either by the SHFE or by trading in the Asian day time trading period (depending on the model employed). This conclusion holds irrespective of the discovery metric employed. It provides important context for the 2012 decision by Hong Kong Exchanges and Clearing (HKEX) to purchase the LME and suggests that the major battle in the coming decade for exchange dominance in copper will be between the SHFE and HKEX-LME in relation to the Chinese market.

The paper is structured as follows. Section 2 provides the copper market context. Section 3 looks at the IS and PT discovery measure. and in section 4 we discuss the use of these measures in the context of non-synchronous trading. Section 5 is devoted to PT discovery estimates and section 6 to IS estimates. In section 7 we consider sub-samples to examine whether the estimated share change over time. Section 8 summarizes results and section 9 concludes.

2. The copper market context

The copper industry experienced a strong cyclical upswing in prices in the first decade of the century that was heavily supported by sustained global industrial expansion. The start of 2003 saw renewed GDP growth in the OECD in conjunction with rapid industrial growth in Asia, particularly China. Low inventory levels and severe supply bottlenecks resulted in a "super cycle" situation in which mine and smelter capacity struggled to keep pace with expanding consumption demand.[2]

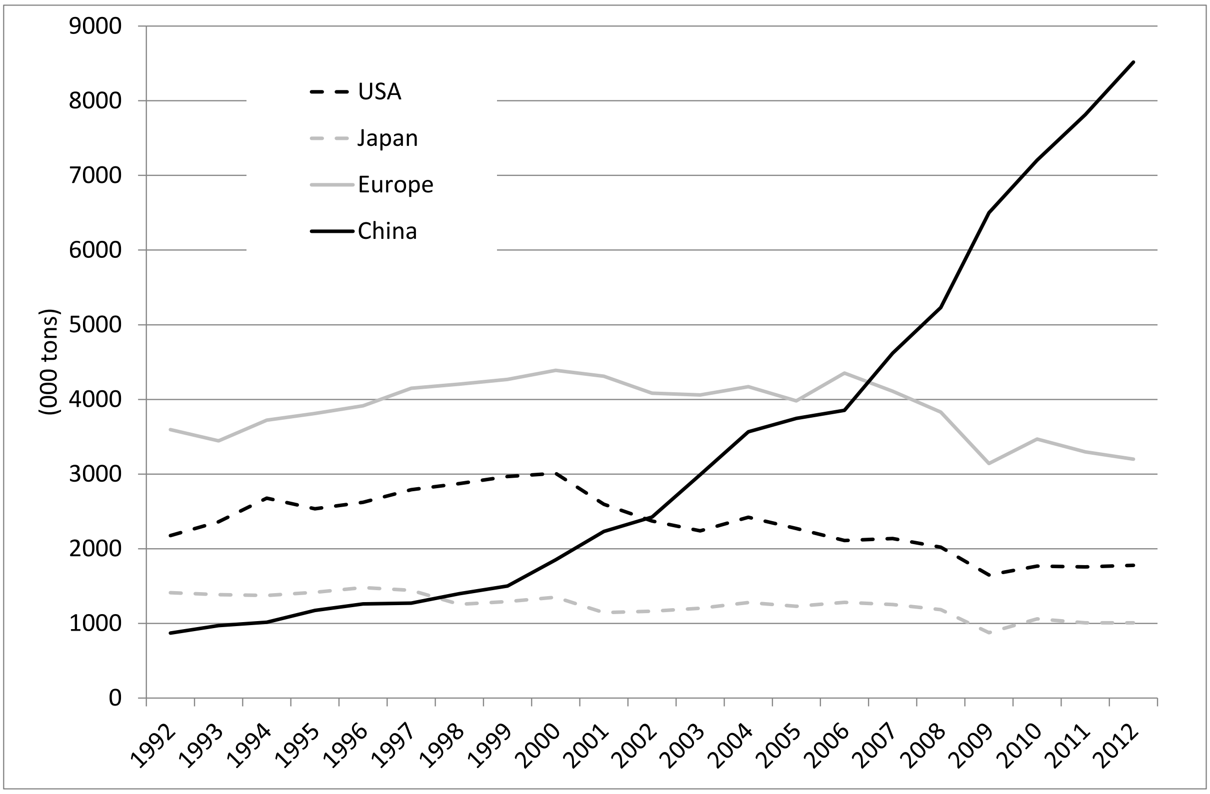

Figure 1: Consumption of refined copper, 1992-2012

Figure 1 shows aggregate copper consumption levels in China, USA, Europe (including Eastern Europe) in thousand (metric) tons. The figures underline the fact that the increased consumption over the last decade was mainly led by China partly at the expense of consumption in Europe and the USA which declined over the same period. Refined production (almost all from imported ores and concentrates) has grown over the same period – see Figure 2. However, the gap between the two, covered by imports of refined metal has increased substantially over the period – from 725,000 tons in 2001 (the start of our sample) to 2.8 million tons in 2012.

[2] See Banks (2011) and the remainder of this Special Report.

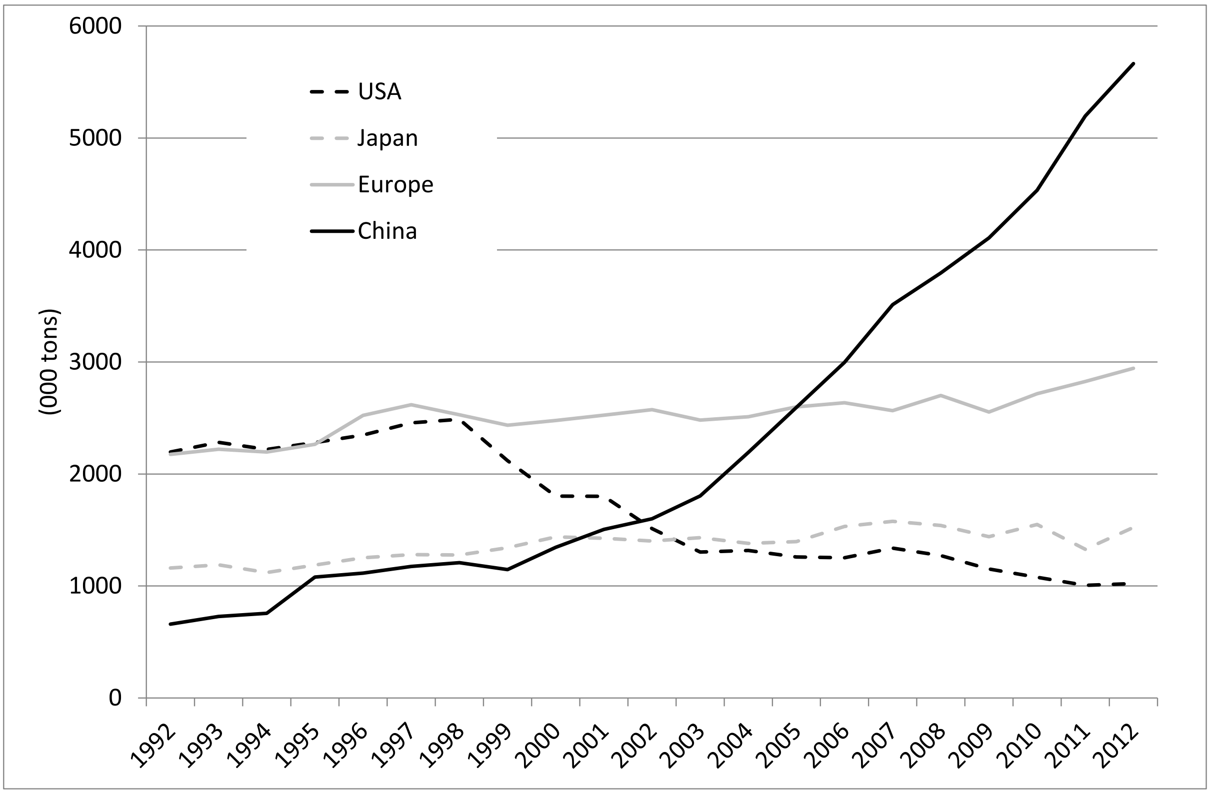

Figure 2: Production of refined copper, 1992-2012

Figures 1 and 2 emphasize the extent to which international commerce in copper has shifted away from Europe and North America, where the LME and Comex operate and have warehouses, to China. Purchases of refined copper for import into China will generally be priced against either the LME or Comex price and sold in China basis the SHFE price. The differential between the two, which have generally been positive, must be such as to ensure the required level of imports. The interplay between the SHFE on the one hand and Comex and the LME on the other is therefore crucial in driving copper commerce.

The differential between internal Chinese copper prices and prices on the world market results from the interplay between a number of factors. Freight charges are important but cannot explain the difference between copper prices in China and in other Asian locations such as Singapore and South Korea. Import duties add between one and two per cent to the internal price. The most complicated factor relates to financing. In the context of Chinese renminbi interest rates which have been higher than U.S. dollar rates throughout the sample we consider and of limited credit availability for private companies, copper imports provide a route to low cost financing. Banks (2011) describes how a Chinese importer can obtain a letter of credit from a western bank by posting 20 per cent margin, sell the copper, which remains in a mainland China bonded warehouse, and invest the remain 80 per cent of the purchase price repaying after 90 or 180 days when the letter of credit expires. This will generally be less expensive than borrowing from a domestic bank. According to industry sources, financing accounts for around one third of the copper imported into China over recent years and, at times, almost all the copper held in bonded warehouses. In the event of a negative differential, copper can be exported from bonded warehouses but this is more complicated than importing. The differential between Chinese and world prices therefore reflects the often rapidly changing Chinese credit market conditions as well as the balance between supply and demand for copper for industrial consumption.

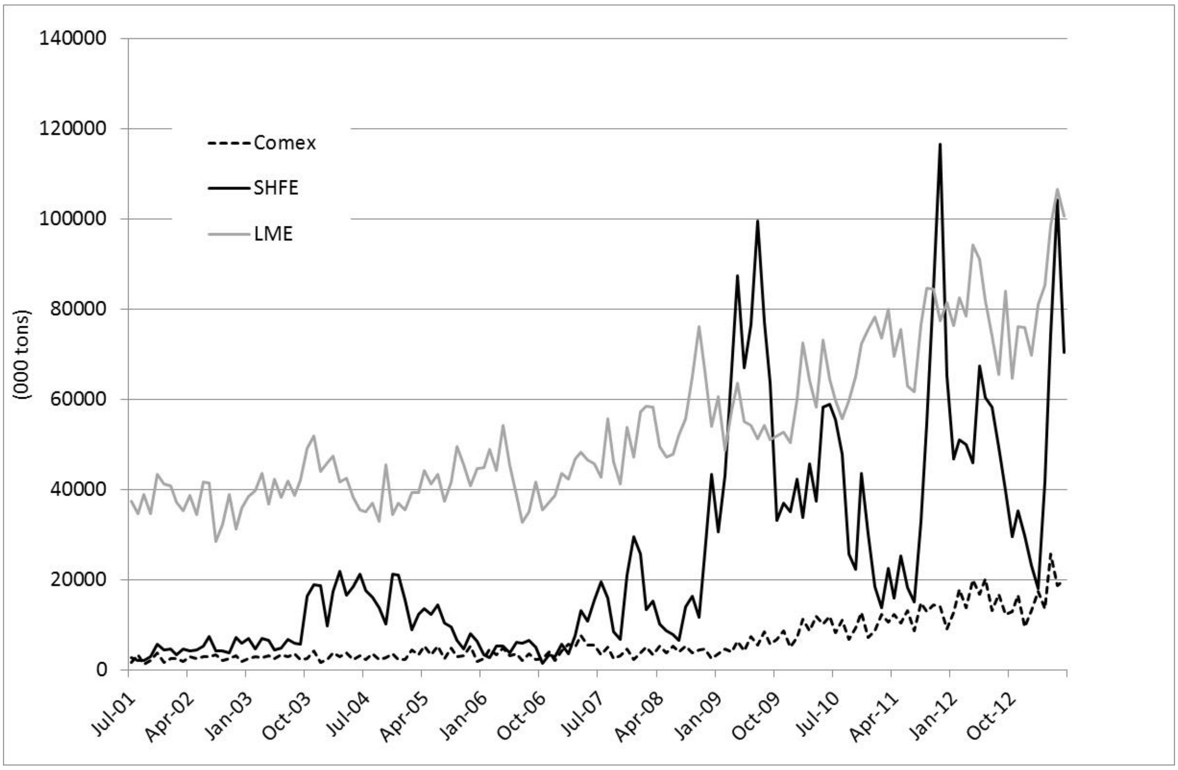

Figure 3: Monthly total trading volumes, July 2001 – June 2013

This increased importance of China in copper commerce has been reflected in an increase in copper trading volumes on the SHFE. Figure 3 shows total monthly volumes (thousand tons) traded in the three markets. Volumes have generally increased over the 2001-2012 period. Figuerola-Ferretti and Gonzalo (2010) use a theoretical model to link the PT price discovery metrics with the relative number of market participants. The greatest growth of volume levels is seen in the Shanghai copper market, and the Figuerola-Ferretti and Gonzalo (2010) model implies that this should play an important role in price discovery in the second half of the sample, at least on the PT measure.

3. Price discovery measures

There are two standard methodologies for measuring price discovery – the Information Share (IS) measure proposed by Hasbrouck (1995) and the Permanent-Transitory (PT) measure proposed by Gonzalo and Granger (1995). The PT and IS measures respond to different questions. The PT measure effectively asks the extent to which the various market prices reflect the long term fundamental and relates to the expected value of the fundamental. We can think of PT as responding to a benchmark question – which is the best price or price average to take as a benchmark for the fundamental price. The IS measure asks about the contribution of the various markets to variation in the fundamental and hence relates to the variance of the fundamental. The fundamental will change in response to the arrival of new information into the market. Relative to the PT measure, the IS measure will give greater weight to the market or markets where most information arrives is impounded into the prices. In general, there is no reason to expect that the market which plays the greatest role in information aggregation will necessarily provide the best price benchmark.

Both methods rely on representation of the prices processes as a Cointegrated Vector Autoregressive (CVAR) system (Johansen, 1991). We follow standard practice in supposing this representation to be in logarithms and, without loss of generality, write the CVAR(*k*) as

$$\begin{aligned}\Delta \ln p_t &= \kappa + \Pi(L)\Delta \ln p_{t-j} + \alpha\beta' \ln p_{t-k} + \varepsilon_t \\ &= \kappa + \sum_{j=1}^{k-1} \Pi_j \Delta \ln p_{t-j} + \alpha\beta' \ln p_{t-k} + \varepsilon_t\end{aligned} \tag{1}$$

Here, p_t is a vector of $m > 1$ prices all of which are integrated of order 1, ε_t is an m-vector of shocks and β is an m by q matrix defining the q cointegrating vectors $(1 \le q < m)$.[3] $\Pi(0)$ is null. We focus on the case in which $q = m - 1$ so that there is a single common trend which we can identify as the underlying fundamental price.

[3] It is more conventional to put the lagged level term at the first lag but this becomes problematic in the analysis of overlapping non-synchronous data – see section 4. The values of α and β in equation (1) are independent of this choice.

In the case we are considering of m-1 cointegrating vectors, the system of equations defined by (1) may be inverted to give the vector moving average (VMA) representation

$$\ln p_t = \ln p_0 + \kappa + \Psi(1)\sum_{j=0}^{t-1}\varepsilon_{t-j} + \Psi^*(L)\varepsilon_t \tag{2}$$

De Jong (2002) shows that we may write $\Psi(1) = \beta_\perp \alpha_\perp{}'$ where $\alpha_\perp$ and $\beta_\perp$ are the orthogonal matrices for α and β respectively and which satisfy $\alpha_\perp \alpha = 0$ and $\beta_\perp \beta = 0$ plus the normalizations $\alpha_\perp \ell_m = \ell_{m-1}$ and $\beta_\perp \ell_m = \ell_{m-1}$ and where ℓ_r is the r-vector of units. The permanent component f of the price may now be identified as

$$(\ln f_t)\,\ell_{m-1} = (\ln f_0)\,\ell_{m-1} + \alpha_\perp{}'\sum_{j=0}^{t-1}\varepsilon_{t-j} = (\ln f_0)\,\ell_{m-1} + \alpha_\perp{}'(\ln p_t - \ln p_0) \tag{3}$$

Because the fundamental price is common to the price in each market, the m-1 rows of $\alpha_\perp{}'$ are identical allowing us to write $\alpha_\perp{}' = \ell_{m-1}\psi'$.

The Gonzalo and Granger (1995) PT discovery measure derives directly from equation (3). They ask how much each market contributes to the fundamental price. This gives

$$PT_j = \frac{\psi_j}{\ell_m{}'\psi} \tag{4}$$

In the case of two markets, equation (4) can be expressed in terms of the α parameters as

$$PT_1 = -\frac{\alpha_2}{\alpha_1 - \alpha_2} \quad PT_2 = \frac{\alpha_1}{\alpha_1 - \alpha_2} \tag{5}$$

If it is established that the two price series are cointegrated, the Granger representation theorem (Engle and Granger, 1987) implies that either $\alpha_1 > 0$ or $\alpha_2 < 0$ or both of these. This allows the possibility of obtaining a negative estimate for α_1 , which would imply a negative discovery share PT_2 or a positive estimate for α_2 , which would imply a negative discovery share PT_1. In such cases, one can impose a zero value on the coefficient implying a unit discovery share for the market in question.

Now consider the general case of m markets. Write $\alpha = \begin{pmatrix} A \\ \alpha_m' \end{pmatrix}$, where A is m-1 by m-1 and similarly $\alpha_\perp = \begin{pmatrix} X & x_m \end{pmatrix}$ where X is also m-1 by m-1 . Then $XA + x_m \alpha_m' = 0$. From the normalization condition $X\ell_{m-1} + x_m = \ell_{m-1}$. Substituting for x_m we get $X = -\left(\ell_{m-1}\alpha_m'\right)\left(A - \ell_{m-1}\alpha_m'\right)$. Hence $X = -\left(\ell_{m-1}\alpha_m'\right)\left(A - \ell_{m-1}\alpha_m'\right)^{-1}$ and

$$\alpha_\perp = \left(\left(-\ell_{m-1}\alpha_m'\right)\left(A - \ell_{m-1}\alpha_m'\right)^{-1} \quad \ell_m - \left(-\ell_{m-1}\alpha_m'\right)\left(A - \ell_{m-1}\alpha_m'\right)^{-1}\ell_m \right) \qquad (6)$$

The PT discovery measures automatically sum to unity and are uniquely determined given the VAR specification (1). However, they rely on accurately determined α coefficients.

As in the bivariate case, there is no guarantee that the share assigned to any market will lie within the unit interval. Consider the jth market. Since we can characterize the cointegrating basis in terms of a set of unit cointegrating vectors (i.e. the log difference between price pairs), we can normalize the β matrix such that $\beta_{ji} = 1$ for each cointegrating vector $i = 1,\ldots,m-1$. A negative share estimate may arise if the corresponding α coefficient $\alpha_{ji} > 0$. This possibility is excluded if the corresponding α coefficients are restricted to zero. However, in the general case, there may not be a unique set of zero restrictions which attain this result. Hence, while it will generally be possible to respecify the VAR to obtain acceptable PT measures, this process involves exercise of judgement which undermines the claim that the PT procedure gives unique and unambiguous discovery share estimates.

The PT measure relies solely on the contribution of each market to the fundamental price f. The IS measure extends this to ask how much each market contributes to the variability of the fundamental. This involves taking into account the variability of the disturbances ε. Consider the effect of shocks to each variable on this fundamental price. From equations (2) and (3), $\Psi(1) = \alpha_\perp' = \psi \ell_{m-1}'$, say, since each row of $\alpha_\perp$ is identical. Let $\Omega = Var(\varepsilon_t)$ which we take to constant over time. In the case in which the shocks to each market are independent and so $\Omega = \begin{pmatrix} \omega_1^2 & & 0 \\ & \ddots & \\ 0 & & \omega_m^2 \end{pmatrix}$, we obtain $Var(\Delta \ln f_t) = \sum_{j=1}^{m} a_j^2 \omega_j^2$. In this case, Hasbrouck (1995) defined the information share IS_j of the jth market as

$$IS_j = \frac{\psi_j^2\omega_j^2}{\sum_{i=1}^{m}\psi_i^2\omega_i^2} = \frac{\psi_j^2\omega_j^2}{\psi'\Omega\psi} \tag{7}$$

Definition (7) can also be employed when Ω is non-diagonal but in this case the supposed shares will not sum to unity. Typically, we find that shocks are positively correlated across markets with the consequence that the summed shares exceed unity. The standard response is to diagonalize Ω through the Cholesky factorization. Define a new set of mutually orthogonal shocks ν_t satisfying $Var(\nu_t) = I_m$. We can write $\varepsilon_t = Q'\nu_t$ where $Q = \begin{pmatrix} q_1' \\ \vdots \\ q_m' \end{pmatrix} = \begin{pmatrix} 1 & & & 0 \\ q_{21} & q_{22} & & \\ \vdots & & \ddots & \\ q_{m1} & q_{m2} & \cdots & q_{mm} \end{pmatrix}$ is lower triangular. It follows that $\Omega = Q'Q$. In this notation, $Var(\Delta \ln f_t) = q'q = \psi'Q'Q\psi$. Under the Cholesky factorization, the information share becomes

$$CIS_j = \frac{q_j^2}{q'q} \tag{8}$$

These shares sum to unity but are dependent on the ordering of the markets. In the case of two markets the two alternative orderings define an upper and a lower bound on the share of each market and it is a common practice to obtain a compromise estimate by averaging these bounds – see Baillie et al. (2002). With three markets, there are six possible orderings and it becomes very difficult to interpret the resulting measures.

There is no way of unambiguously resolving the error correlations. This motivates decomposition based on posited factor structures as in Lien and Shrestha (2009, 2014). Posit $g \le m$ mutually orthogonal factors $z_1,\ldots,z_g$ with $Var(z_t) = \Theta = diag(\theta)$ and such that $\varepsilon_t = Gz_t$. Then $Var(\varepsilon_t) = \Omega = G\Theta G'$ and $Var(\Delta \ln f_t) = q'q = \psi'G\Theta G'\psi$. The information share of factor *j* is given by equation (8) with $q = F'\psi = \Theta^{1/2}G'\psi$. The information shares are not unique since they depend on the factor structure adopted. Principal components analysis (PCA) is a commonly used factor structure which fits into this category. Applied to the correlation matrix of the shocks

$R = \hat{\omega}^{-½}\Omega\hat{\omega}^{-½}$ where $\hat{\omega} = \begin{pmatrix} \omega_1^2 & & 0 \\ & \ddots & \\ 0 & & \omega_m^2 \end{pmatrix}$, PCA chooses the factors such that the first factor (component) $z_{1t} = \frac{\gamma_1{'}\varepsilon_t}{\omega_1}$ is the maximum variance linear combination of the disturbances subject to the Euclidean normalization constraint $\gamma_1{'}\gamma_1 = 1$. Subsequent components maximize the variances of the shocks orthogonalized to the earlier components. The eigenvalue decomposition of the correlation matrix is $R = \Gamma\Lambda\Gamma{'}$ and hence $\Theta = \Lambda = diag(\lambda)$, the matrix with the eigenvalues of R along the diagonal, $G = \hat{\omega}^{½}\Gamma$ and

$$q = F{'}\psi = \Lambda^{½}\Gamma{'}\hat{\omega}^{½}\psi \qquad (9)$$

The factor information shares are given by equation (8) as previously. They will be independent of the order of the variables in the VAR. The procedure is straightforward and unambiguous but the factors which perform this variance reduction will not necessarily have a natural interpretation.

The factor IS model yields information shares be interpreted in terms of factors and not markets. In the context of the PCA factor decomposition this requires that the principal components should be interpretable. In practice, the first principal component will invariably be a weighted average giving approximately equal weights to all *m* prices. We can interpret this as an overall market price, similar to the permanent component identified by PT models. Lower order components may be less interpretable but will often correspond to relativities between prices. Factor IS shares therefore relate only indirectly to the question of which markets contribute most to discovery.

A standard procedure is factor analysis, for example in educational psychology, is to rotate the principal components – see Morrison (1976, chapter 9). Let *T* be any orthogonal matrix such that $T{'}T = I_m$. Then $F^* = FT{'}$ is a rotated representation of the factor structure. The rotation matrix may be chosen to increase the interpretability of the resulting factors. Lien and Shrestha (2009) offer a factorization which yields what they call the Modified Information Share (MIS) – see also

Lien and Shrestha (2014). The MIS may be interpreted as a rotation of the PCA factor shares. It results from setting $\Theta = I_m$ and $G = \hat{\omega}^{1/2}\Gamma\Lambda^{1/2}\Gamma'$. In terms of the PCA factor structure, this gives[4].

$$q^* = F^*{}'\psi = \Gamma F'\psi \tag{10}$$

Since $\Gamma'\Gamma = I_m$, this amounts to an orthogonal rotation of the original principal components. The rotation adopted by LS weights each of the PCA factors in the proportion that they contribute to each market price. This allows Lien and Shrestha (2009) to interpret the MIS discovery shares in terms of markets which, in this specific case, correspond to the appropriately weighted factors.

4. Data

We analyze daily prices in three markets located in different continents: official LME Settlement (second morning ring) and unofficial (second afternoon ring) prices and Comex and SHFE closing prices. The LME Settlement prices are matched against Comex and SHFE front (first month) contracts while the LME three month prices are matched against Comex and SHFE fourth month prices.[5] The sample is 2 July 2001 to 28 June 2013. Comex closing prices are determined at 13:00 local time. The SHFE close is 15:00 local time, equivalent to 02:00 in New York. Official LME copper prices are determined at 12:35 local time equivalent to 07:30 in New York while the unofficial prices are declared at 16:15, equivalent to 10:15 in New York.[6] The resulting price

---

[4] In the case of two markets with $R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ one can evaluate $\Lambda = \begin{pmatrix} 1+\rho & 0 \\ 0 & 1-\rho \end{pmatrix}$ and $\Gamma = \begin{pmatrix} ½\sqrt{2} & -½\sqrt{2} \\ ½\sqrt{2} & ½\sqrt{2} \end{pmatrix}$. The PCA factor structure gives $F' = \begin{pmatrix} \frac{\sqrt{2}}{2}\omega_1\sqrt{1+\rho} & \frac{\sqrt{2}}{2}\omega_2\sqrt{1+\rho} \\ -\frac{\sqrt{2}}{2}\omega_1\sqrt{1-\rho} & \frac{\sqrt{2}}{2}\omega_2\sqrt{1-\rho} \end{pmatrix}$. Lien and Shrestha (2009) set $\Theta = I_m$ and $G = \hat{\omega}^{1/2}\Gamma\Lambda^{1/2}\Gamma'$. In terms of the PCA factor structure, this gives $q^* = F^*{}'\psi = \Gamma F'\psi$. Performing the multiplication in the case of two markets, one obtains $F^*{}' = \begin{pmatrix} ½\omega_1\left[\sqrt{1+\rho}+\sqrt{1-\rho}\right] & ½\omega_2\left[\sqrt{1+\rho}-\sqrt{1-\rho}\right] \\ ½\omega_1\left[\sqrt{1+\rho}-\sqrt{1-\rho}\right] & ½\omega_2\left[\sqrt{1+\rho}+\sqrt{1-\rho}\right] \end{pmatrix}$ which is the result they report.

[5] SHFE settlement prices are daily value-weighted average prices (VWAPs). For this reason we use the closing and not the Settlement prices. Data sources: Comex: Norma's Historical Data, LME: LME, SHFE: Bloomberg.

[6] Conversions of LME and SHFE prices to New York time will vary at certain times of year due to non-synchronized daylight saving arrangements.

change observations will be overlapping as shown in Figure 4. (LMES refers to the LME Settlement price and LMEU to the unofficial price).

The data present two other issues. The first relates to days in which only one or two of the three markets traded. One possibility would be to eliminate such days from the sample but this would be complicated in the context of non-synchronous trading. The alternative, which we have followed, is to maintain all days in the sample but to infill prices on non-trading days with the most recent but stale price from the same exchange.

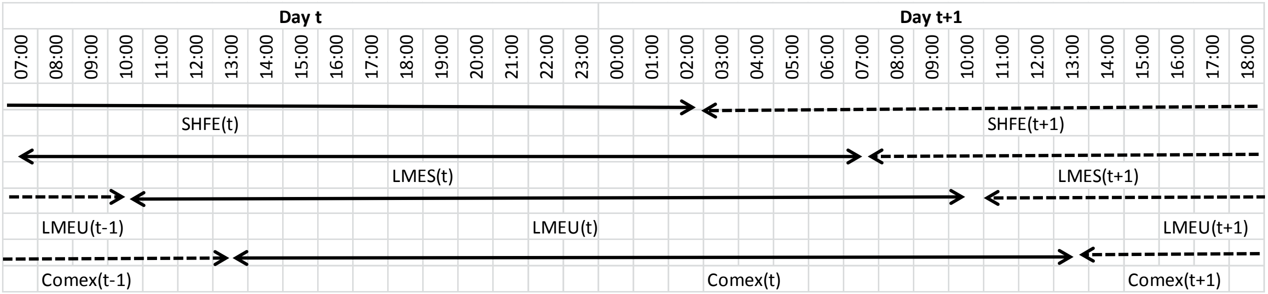

Figure 4: Overlapping price reference times

The second problem relates to rolling at contract expiration dates. This problem arises in conjunction with the Comex and SHFE prices but not with LME prices since the LME contract structure involves each trading day being the prompt date for a contract which expires on that date. It would therefore be inappropriate to roll these contracts. A symmetric treatment of Comex and SHFE prices therefore requires that these should also not be rolled. The VAR specification (1) argues in the same direction since roll adjustments imply that rolled price changes are not equal to the differences between price levels when these prices relate to different contracts. The consequence is that the Johansen VAR specification (1) will involve additional nuisance terms arising from monthly roll returns. For these reasons, we choose to treat both the Comex and SHFE prices series as continuous futures.

Stationarity tests are reported in Table 1. The logarithms of all four prices are I(1). A Johansen (1989) cointegration test establishes that there are three cointegrating vectors both for the front and the deferred prices. This implies that all four pieces are cointegrated at each horizon.[7] This permits us to select an arbitrary cointegrating basis. We choose the logarithmic differences between the LME Settlement prices and respectively the LME unofficial prices, Comex prices and

[7] We do not consider cointegration between front and deferred prices in this paper.

SHFE and LME Settlement prices. All three differences are I(0) justifying the imposition of unit cointegrating vectors. (Other combinations of these four level variables would have proved equally valid).

Table 1 ADF tests					
		SHFE (1)	LMES (2)	LMEU (3)	Comex (4)
Front prices	$\ln p_{jt}$	ADF(18) -1.60	ADF(11) -1.61	ADF(11) -1.61	ADF(11) -1.62
	$\Delta\ln p_{jt}$	ADF(17) -10.2	ADF(11) -13.2	ADF(10) -13.8	ADF(9) -15.3
	$\ln p_{jt} - \ln p_{2t}$	ADF(14) -4.35	-	ADF(13) -10.3	ADF(9) -6.70
3 month deferred prices	$\ln p_{jt}$	ADF(8) -1.50	ADF(12) -1.59	ADF(4) -1.56	ADF(10) -1.58
	$\Delta\ln p_{jt}$	ADF(9) -15.3	ADF(10) -13.9	ADF(3) -27.8	ADF(9) -15.4
	$\ln p_{jt} - \ln p_{2t}$	ADF(10) -4.54	-	ADF(5) -18.1	ADF(8) -12.1
The lag length *k* is chosen using the Akaike Information Criterion (AIC). Sample 2 July 2001 to 28 June 2013. Lag lengths were selected using the Akaike Information Criterion (AIC). Critical values: 5% -2.86, 1% -3.44.					

5. PT price discovery estimates

The PT measures are calculated using equation (6) from a VAR written in the form (1). However, this is complicated by the non-synchronous (overlapping) structure of the data we analyze. Standard VAR models, as that specified as equations (1), condition on past variables. Order the markets in order of closing, SHFE (1), LMES (2), LMEU (3) and Comex (4). It is clear from Figure 4 that the time period covered by $\Delta\ln p_{1t}$ over laps $\Delta\ln p_{2,t-1}$ and $\Delta\ln p_{3,t-1}$. Similarly, $\Delta\ln p_{2t}$ overlaps $\Delta\ln p_{3,t-1}$ and $\Delta\ln p_{4,t-1}$ but not $\Delta\ln p_{1,t-1}$. Similarly arguments apply for $\Delta\ln p_{3t}$ and $\Delta\ln p_{4t}$. The standard VAR philosophy requires that all regressor variables be predetermined. This entails deletion of those regressors at the first lag which overlap one of the dependent variables. If one wishes to estimate a standard VAR(*k*), equations (1) therefore need to be respecified as

$$
\begin{aligned}
\Delta \ln p_{1t} &= \kappa_1 + \pi_{111}\Delta \ln p_{1,t-1} + \sum_{j=2}^{k-1}\pi_{1j}{}'\Delta \ln p_{t-j} + \alpha_1{}'\beta{}'\ln p_{t-k} + \varepsilon_{1t} \\
\Delta \ln p_{2t} &= \kappa_2 + \sum_{i=1}^{2}\pi_{2i1}\Delta \ln p_{i,t-1} + \sum_{j=2}^{k-1}\pi_{2j}{}'\Delta \ln p_{t-j} + \alpha_2{}'\beta{}'\ln p_{t-k} + \varepsilon_{2t} \\
\Delta \ln p_{3t} &= \kappa_3 + \sum_{i=1}^{3}\pi_{3i1}\Delta \ln p_{i,t-1} + \sum_{j=2}^{k-1}\pi_{3j}{}'\Delta \ln p_{t-j} + \alpha_3{}'\beta{}'\ln p_{t-k} + \varepsilon_{3t} \\
\Delta \ln p_{4t} &= \kappa_3 + \sum_{i=1}^{4}\pi_{4i1}\Delta \ln p_{i,t-1} + \sum_{j=2}^{k-1}\pi_{4j}{}'\Delta \ln p_{t-j} + \alpha_4{}'\beta{}'\ln p_{t-k} + \varepsilon_{4t}
\end{aligned}
\qquad (11)
$$

The system is estimable by Ordinary Least Squares (OLS) with the lag length k determined by minimization of the Akaike Information Criterion (AIC).

Table 2
Estimated market discovery shares using the PT decomposition

	SHFE	LME Settlement	LME Unofficial	Comex
Front prices	35.0%	0.0%	0.0%	65.0%
Three month deferred prices	41.5%	0.0%	0.0%	58.5%

The PT shares are calculated using equation (6) from the estimated α coefficients from the VAR (11) estimated by OLS over the sample 2 July 2001 to 28 June 2013. α coefficient have been set to zero where the unrestricted estimates gave negative PT estimates. Equation estimates are available on request.

Unrestricted estimation of that system of equations (11) yields PT shares outside the unit interval. Hence, as outlined above, the α coefficients responsible for this result were set to zero. The resulting estimated PT discovery shares are reported in Table 2 for both the front and the deferred prices. In both case, the restrictions imply a zero share for the two LME prices with discovery divided between Comex and the SHFE, with Comex playing the larger role.

A high estimated PT share results from low values of the α error correction coefficients since the price appears uninfluenced by shocks to the remaining prices. Conversely, high estimated values for the α error correction coefficients will imply a low PT share since this price is seen as adapting to shocks in other prices. In practice, these inferences can prove problematic since small estimated α coefficients are also likely to be poorly determined. Inferences may turn out not to be robust because they depend on the least precisely estimated coefficients in the system. This feature of the results which give pre-eminence to Comex and the SHFE, carries through to the IS estimates based on the VAR (11) which rely on the same estimated α coefficients.

An alternative approach to the non-synchronous trading problem is to specify a recursive Structural VAR (SVAR) with a recursive structure reflecting the order of trading. The recursive SVAR(*k*) specification is

$$\begin{aligned}
\Delta \ln p_{1t} &= \kappa_1 + \sum_{j=1}^{k-1} \pi_{1j}'\Delta \ln p_{t-j} + \alpha_1'\beta'\ln p_{t-k} + \varepsilon_{1t} \\
\Delta \ln p_{2t} &= \kappa_2 + \pi_{201}\Delta \ln p_{1t} + \sum_{j=1}^{k-1} \pi_{2j}'\Delta \ln p_{t-j} + \alpha_2'\beta'\ln p_{t-k} + \varepsilon_{2t} \\
\Delta \ln p_{3t} &= \kappa_3 + \sum_{i=1}^{2} \pi_{30i}\Delta \ln p_{1t} + \sum_{j=1}^{k-1} \pi_{3j}'\Delta \ln p_{t-j} + \alpha_3'\beta'\ln p_{t-k} + \varepsilon_{3t} \\
\Delta \ln p_{4t} &= \kappa_4 + \sum_{i=1}^{3} \pi_{40i}\Delta \ln p_{1t} + \sum_{j=1}^{k-1} \pi_{4j}'\Delta \ln p_{t-j} + \alpha_4'\beta'\ln p_{t-k} + \varepsilon_{4t}
\end{aligned} \tag{12}$$

The model may be written more compactly as

$$\Delta \ln p_t = \kappa + \Pi(L)\Delta \ln p_t + \alpha\beta'\ln p_{t-k} + \nu_t$$

with $\Pi(0) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \pi_{201} & 0 & 0 & 0 \\ \pi_{301} & \pi_{302} & 0 & 0 \\ \pi_{401} & \pi_{402} & \pi_{403} & 0 \end{pmatrix}$. Equations (10) reflect the fact that the SHFE closing price is already determined at the time of the second LME ring which determines the official LME prices and that both this and the official LME prices are known in Comex closing period.

This specification has some advantages over the standard VAR specification (12) but also changes the interpretation of the estimated coefficients. OLS estimation of (10) orthogonalizes the residuals ν_{2t} with respect to $\Delta \ln p_{1t}$ and ν_{3t} with respect to both $\Delta \ln p_{1t}$ and $\Delta \ln p_{2t}$. Consequently OLS estimation imposes the condition $E[\nu_t \nu_t] = diag(\omega)$.[8]

[8] Hua et al. (2010), who recognize the problem of non-synchronous trading, follow a procedure which is a hybrid between equations (1) and (12). They forward date the SHFE price, here p_1, by one day in a VAR of form (1) but with the lagged level terms at the first lag. In terms of the notation in equation (12), they set $\pi_{201} = 0$ but leave π_{301} and π_{302} unrestricted. The result is to orthogonalize the disturbances on the LME

This restriction changes the interpretation of the estimated shares which now relate to time slots and not markets. In our implementation, the disturbance ε_{4t} relates to the Comex dependent variable $\Delta \ln p_{4t}$ but this is orthogonalized with respect to the disturbance ε_{3t} relating to the change in the unofficial LME price $\Delta \ln p_{3t}$. Interpreting the disturbance in terms of the arrival of information into the market, information which arrives prior to the 10:15 EST declaration of the LME unofficial price will be included in ε_{3t} irrespective of whether that information first affected prices in Comex or on the LME. The disturbance ε_{4t} therefore relates to information arriving after 10:15 EST. This motivates division of the trading day into four time slots: London a.m., from the closure of the SHFE to the end of the second LME morning ring, London p.m., from the second morning ring to the second afternoon ring, New York a.m. from the second LME ring to the Comex closure, and the Shanghai day, from the Comex to the SHFE closure.

Table 3
Estimated time slot discovery shares using the PT decomposition

	London a.m. 03:01 – 07:35	London p.m. 07:36 – 10:15	New York a.m. 10:16 – 13:00	Shanghai day 13:01 – 03:00
Front prices	9.7%	35.2%	25.7%	29.4%
Deferred prices	19.4%	7.4%	32.1%	41.1%

Time slots are EST with LME and SHFE closing times converted to EST for dates on which there is no daylight saving. The PT shares are calculated using equation (6) from the estimated α coefficients from the VAR (12) estimated by OLS over the sample 2 July 2001 to 28 June 2013. Equation estimates are available on request.

The estimated α coefficients from the recursive SVAR (12) are much more precisely determined than those from the standard VAR. Although it remains possible to obtain PT shares outside the unit interval, this did not happen in our full sample estimates. Estimated shares are given in Table 3. They show around 30% of the discovery taking place in the New York morning after the LME rings and a further 30%-40% in the Asian day trading slot leaving around 30%-40% for the two London slots. However, absent transactions data, it is not possible to attribute responsibility to any particular market. During the time that the SHFE is actively trading, for example, it is also possible to trade on the electronic LME Select system and any such transactions may impact SHFE prices.

---

and Comex equations respectively with that on the SHFE equation but to allow a non-zero correlation between the LME and Comex disturbances.

Aggregating the shares for the two London time slots, the time zone discovery shares summarized in Table 3 divide discovery approximately equally across the time zones associated with the three continents. No single time zone is seen as dominant. The estimates fail to offer support for the prevalent practice of LME Settlement prices as the most reliable copper benchmark.

6. IS price discovery estimates

IS discovery shares may be computed either from the standard overlapping observation VAR defined by equation (11) or from the recursive SVAR defined by equations (10). The second approach is the most straightforward since the error variance matrix Ω is diagonal implying unambiguous IS shares which automatically sum to unity.

Table 4
Estimated time slot discovery shares using the IS decomposition

	London a.m. 03:01 – 07:35	London p.m. 07:36 – 10:15	New York a.m. 10:16 – 13:00	Shanghai day 13:01 – 03:00
Front prices	6.0%	0.2%	44.8%	49.0%
Deferred prices	12.0%	3.0%	33.7%	51.3%

Time slots are EST with LME and SHFE closing times converted to EST for dates on which there is no daylight saving. The IS shares are calculated using equation (7) from the estimated α coefficients from the VAR (12) estimated by OLS over the sample 2 July 2001 to 28 June 2013. Equation estimates are available on request.

Table 4 reports results based on the recursive VAR (12). These are directly comparable in terms of interpretation with the PT estimates reported in Table 3. In contrast with the PT results, which divided discovery fairly equally between Europe, North America and Asia, the IS estimates attribute discovery primarily to the New York morning and Shanghai time slots. The London morning and afternoon slots are seen as much less important. The relative unimportance of the two London slots in these estimates suggests that little new information arrives in the market in European trading time.

We now turn to the Cholesky-IS estimates defined by equation (8) in conjunction with the estimates of the overlapping VAR as specified in equations (11). As noted, these estimates depend on the ordering of the variables. With four prices, we have 24 possible orderings. Table 5 reports the minimum, mean and maximum values of these estimates for each market. These estimates underline the limitations of the IS procedure when employed with non-orthogonal

VAR more than they inform about the price discovery process. If one follows the common procedure of looking at the shares averaged over orderings, one concludes here, as in other studies, that all markets contribute to discovery although, on this criterion, Comex appears substantially more important than the SHFE and LME.

Table 5 **Estimated market discovery shares using the Cholesky-IS decomposition**					
		SHFE	LME Settlement	LME Unofficial	Comex
Front prices	minimum	5.5%	0.0%	0.0%	38.9%
	average	13.7%	15.7%	15.8%	54.8%
	maximum	37.2%	54.5%	54.6%	93.5%
Three month deferred prices	minimum	7.0%	0.0%	0.0%	23.8%
	average	19.3%	19.7%	15.1%	45.9%
	maximum	49.7%	62.8%	50.0%	91.4%
The Cholesky-IS shares are calculated using equation (8) from the estimated α coefficients and error variance matrix Ω from the VAR (11) estimated by OLS over the sample 2 July 2001 to 28 June 2013. We report the minimum, mean and maximum shares over the 24 possible orderings of the four markets.					

The Cholesky-IS procedure has the advantage that IS discovery shares are uniquely determined. Nevertheless, orthogonalization does entail an implicit factor model is which innovation v_{jt} represents the information arriving on day *t* in the time slot between the closure of market *j*-1 and the closure of market *j*. Consider, for example, the 2¾ hour time slot between the 07:30 EST determination of the LME settlement prices and the 10:15 EST determination of LME unofficial prices. Comex is already actively trading over the final two hours of this period and hence information arriving in the markets, accounted for in the model by the innovation v_{3t}, is impounded in both the Comex and the LME prices. The Cholesky-IS procedure attributes this information to whichever of Comex and the LME appears earlier in the variable ordering. That attribution in necessarily arbitrary and an average of arbitrary statistics remains arbitrary. Furthermore, these conclusions are subject to the qualification that they rest on poorly determined α error correction coefficients. We conclude that it is difficult to judge the relative importance of different markets in the absence of high frequency data covering time periods in which at least two markets are open.

Finally, we report the principal component factor IS estimates. These have the merit of explicit adoption of a factor structure. The factor loadings, normalized such that the absolute values of the loadings sum to unity, are shown in the final four columns of Table 6. The leading principal component is close to a simple average of the innovations in each market. The second component is a contrast between the SHFE and the non-Chinese markets. The third component for the front contract and the fourth component for the deferred contract are contrasts between LME and Comex prices. The fourth component is irrelevant in the front decomposition while that for the deferred decomposition is difficult to interpret.

The estimated discovery shares differ between the two contracts. The first principal component is associated with almost 80% of the discovery for the front contract and almost 90% at three months. The PCS model does not attribute this dominant discovery share across markets. Little remains for the other factors which resemble sort term noise although the Comex-LME differential, reflected in the third component, has some importance for front prices. The SHFE differential, reflected in the more important second component, does not appear to contribute to price discovery. As was the case with the Cholesky-IS estimates, these conclusions are subject to the qualification that they rest on poorly determined α error correction coefficients.

Table 6
Principal component IS discovery shares and factor loadings

	Component	Variance share	Discovery share	SHFE	LME Settlement	LME Unofficial	Comex
Front	1	75.0%	79.5%	0.2030	0.2794	0.2795	0.2381
	2	15.7%	0.6%	0.5577	-0.0955	-0.0950	-0.2518
	3	9.3%	19.9%	0.1172	-0.2266	-0.2256	0.4306
	4	0.0%	0.0%	0.0006	0.4990	-0.4998	0.0006
3 month deferred	1	70.5%	89.0%	0.2140	0.2709	0.2519	0.2631
	2	15.1%	0.1%	0.5216	0.0211	-0.2687	-0.1887
	3	7.5%	3.0%	0.1206	-0.2134	0.4024	-0.2636
	4	5.9%	7.9%	-0.1393	0.4496	0.0314	-0.3797

The IS shares are calculated using equations (8) and (9) from the estimated α coefficients from the VAR (12). The factor loadings are the principal components of the VAR residuals. Equation estimates are available on request.

As discussed in section 3, the Lien and Shrestha (2009) Modified Information Share (MIS) model redistributes the factor loadings across markets to generate a market interpretation of the factor loadings. In view of the dominance of the initial overall market factor and the fact that all four prices contribute to that factor (see Table 6), the MIS shares also attribute share to each market. They are reported in Table 7. They show Comex accounting for over one half of the price discovery for both the front and deferred contracts with the remainder divided between the

LME and the SHFE. The SHFE is slightly more important and Comex slightly less important for the deferred contract relative to the front contract. . As was the case with the Cholesky and PCA IS estimates, these conclusions are subject to the qualification that they rest on poorly determined α error correction coefficients.

Table 7
Estimated market discovery shares using the MIS decomposition

	SHFE	LME Settlement	LME Unofficial	Comex
Front prices	14.6%	10.4%	10.5%	64.5%
Three month deferred prices	20.5%	13.7%	10.0%	55.8%

The MIS shares are calculated using equations (8) and (10) from the estimated α coefficients from the VAR (12) estimated by OLS over the sample 2 July 2001 to 28 June 2013. Equation estimates are available on request.

7. Results for sub-samples

Section 2 documented the growth in importance of China as a copper consumer and the associated growth in SHFE trading volumes. This suggests that there may also have been an increase in the role of the Chinese market in price discovery. We investigate this by splitting our twelve year sample into three equal four year sub-samples - 2 July 2001 to 30 June 2005, 1 July 2005 to 30 June 2009, which includes the financial crisis period and the post-crisis period 1 July 2009 to 28 June 2013. We employ both the IS decompositions in relation to the recursive SVAR model defined by equations (12). (We do not pursue the alternative of the overlapping observation VAR defined by equations (11) because of imprecise determination the α error correction coefficients).

The results from applying the PT decomposition are reported in Table 8 and those from the IS decomposition in Table 9. As emphasized in sections 5 and 6, these discovery shares relate in the first instance to time slots and not markets. The differences across sub-samples are relatively minor and may well be due to sampling error. Importantly, the estimates demonstrate the robustness of the conclusions reached in sections 5 and 6 in relation to the importance of the Chinese daytime trading slot.

Table 8 **Estimated time slot discovery shares using the PT decomposition**					
		London a.m. 03:01 – 07:35	London p.m. 07:36 – 10:15	New York a.m. 10:16 – 13:00	Shanghai day 13:01 – 03:00
Front prices	2001-05	8.2%	48.5%	16.1%	27.1%
	2005-09	6.9%	58.0%	20.0%	15.1%
	2009-13*	26.2%	14.9%	24.3%	34.6%
	2001-13	9.7%	35.2%	25.7%	29.4%
Deferred prices	2001-05	22.9%	8.1%	26.0%	43.0%
	2005-09	16.6%	19.1%	31.4%	32.9%
	2009-13	34.5%	1.3%	31.7%	32.5%
	2001-13	19.4%	7.4%	32.1%	41.1%

Time slots are EST with LME and SHFE closing times converted to EST for dates on which there is no daylight saving. The PT shares are calculated using equation (6) from the estimated α coefficients from the recursive SVAR (12) estimated by OLS over the samples 2 July 2001 to 30 June 2005 (rows 1 and 5), 1 July 2005 to 30 June 2009 (rows 2 and 6), 1 July 2009 to 28 June 2013 (rows 3 and 7) and 2 July 2001 to 28 June 2013 (rows 4 and 8). Equation estimates are available on request.

* The front contract estimates for 2009-13 impose $\alpha_{12} = 0$.

Table 9 **Estimated time slot discovery shares using the IS decomposition**					
		London a.m. 03:01 – 07:35	London p.m. 07:36 – 10:15	New York a.m. 10:16 – 13:00	Shanghai day 13:01 – 03:00
Front prices	2001-05	6.0%	3.2%	34.8%	56.0%
	2005-09	7.3%	0.2%	65.5%	27.0%
	2009-13*	26.0%	0.0%	28.9%	45.0%
	2001-13	6.0%	0.2%	44.8%	49.0%
Deferred prices	2001-05	16.0%	2.9%	24.8%	56.3%
	2005-09	13.6%	11.2%	30.7%	44.5%
	2009-13	31.2%	0.3%	33.4%	35.1%
	2001-13	12.0%	3.0%	33.7%	51.3%

Time slots are EST with LME and SHFE closing times converted to EST for dates on which there is no daylight saving. The IS shares are calculated using equation (7) from the estimated α coefficients from the recursive SVAR (12) estimated by OLS over the samples 2 July 2001 to 30 June 2005 (rows 1 and 5), 1 July 2005 to 30 June 2009 (rows 2 and 6), 1 July 2009 to 28 June 2013 (rows 3 and 7) and 2 July 2001 to 28 June 2013 (rows 4 and 8). Equation estimates are available on request.

* The front contract estimates for 2009-13 impose $\alpha_{12} = 0$.

8. Summary of results

Methodologically, the IS and PT discovery measures answer different questions and these differences become clearer once non-synchronous trading is taken into account. LME prices or London daytime trading, depending on the VAR specification, generally show up as less important using the IS measure probably reflecting the greater importance of the U.S. and Chinese markets in generating market-relevant information.

Two alternative VAR specifications are available to account for non-synchronous trading. The first, which maintains an overlapping error structure, failed to yield well determined error correction (α) coefficients in our estimates, and this qualifies the conclusions that can be drawn from these estimates. The alternative recursive SVAR approach gives better determined estimates. However, the resulting discovery shares relate, at least in the first instance, to trading time slots and not to markets.

Table 10
Methodological summary

	PT discovery measure	IS discovery measure
Standard VAR	Poorly determined error correction (α) coefficients result in unreliable or nonsensical estimated shares – see Table 2.	Correlated error terms oblige use of an implicit or explicit factor model. The resulting discovery shares relate to factors and not, in the first instance, markets – see Tables 5 and 6.
Recursive SVAR	The error correction (α) coefficients are well determined but estimated shares relate to time slots and not markets – see Table 3.	As with the PT measure, estimated shares relate to time slots and not markets. The IS shares take into account the error terms as well as the α coefficients – see Table 4.

The table summarizes methodological conclusions. Greater detail is provided in the text.

The ambiguity in the IS discovery measures, first noted by Hasbrouck (1995), points to a fundamental identification problem and cannot be dismissed as simply a nuisance. This issue might be taken as favouring the PT measurement approach where this problem does not arise. The standardly used Cholesky factor structure gives rise to very wide bounds on the discovery shares. Since these estimates depend on essentially arbitrary orderings, the average estimates are also arbitrary. A principal component factor structure provides unambiguous discovery shares but these relate to the factors and not to markets (or time slots). The Lien and Shrestha Modified Information Share (MIS) measure rotates the factor loadings such that the resulting shares again relate to markets.

These methodological conclusions are summarized in Table 10 and the substantive conclusion in Table 11.

Table 11 Summary of substantive conclusions		
	PT discovery measure	IS discovery measure
Standard VAR	Discovery divides between Comex and the SHFE with Comex seen as the more important. The LME does not contribute to discovery – see Table 2. These results may reflect poorly determined α coefficients.	The range of estimated share is too wide to be useful Using the Cholesky decomposition (Table 5). The PCA attributed 80%-90% of discovery to an overall market factor (Table 6). The Modifies IS measure gives weights to all three markets but sees Comex as responsible for over 50% of discovery (Table 7).
Recursive SVAR	Discovery divides between the Shanghai day and the time slot defined by the New York day and London afternoon. The London morning trading slot is less important – see Table 3. There is no strong evidence of any change over time in these rankings (Table 8).	Discovery divides between the Shanghai and New York trading time slots. The London trading slot is seen as much less important – see Table 4. There is no strong evidence of any change over time in these rankings (Table 9).
The table summarizes substantive conclusions. Greater detail is provided in the text.		

9. Conclusions

Over the past two decades, China has become the most important world theatre for international commerce in copper. China is now the dominant consumer of refined copper consuming over four times as much as the United States and two and a half times as much as Europe. China needs to import almost all her requirements of copper, either as ore and concentrate for domestic refining, or as refined copper. International price formation in copper is therefore driven by China's import requirements.

The London Metal Exchange (LME) has been the most important copper futures market, at least outside the USA, for over 50 years. Comex plays a similar role on the North American market. It is to be expected that, with the shift in copper consumption away from Europe and North America and towards China, futures trading in copper would also move to China. The volume of trading of copper futures on the Shanghai Futures Exchange (SHFE) has grown dramatically in response to the changed geographical distribution of the underlying physical market. At the same time, the Chinese copper market is imperfectly integrated with the international market and significant price differences can arise between copper inside and outside China.

Price discovery is the process by which information on current and future developments in production and consumption become impounded in the futures price. A natural question is therefore whether the increased trading volumes on the SHFE are matched by an increased role in copper price discovery. We use end-of-day closing data on front and deferred contracts to infer discovery shares. The resulting data reflect non-synchronous trading and hence overlapping observation periods. This significantly complicates the analysis relative to earlier discussions. It turns out that this apparently simple question cannot be answered in a simple manner using data of this sort. Interesting conclusions, both methodological and substantive, nevertheless emerge.

Two methodologies exist for quantifying the role of competing markets in price discovery. The IS metric introduced by Hasbrouck (1995) and the Gonzalo and Granger (1995) PT discovery measure. The PT measure is based on the extent to which different prices contribute to the expected value of the underlying price fundamental. It relates to the issue of which market best provides a price benchmark. The IS measure looks at the sources of variability of the price fundamental. This relates to the information arrival process and asks in which market most new market-relevant information is impounded into prices. These are different questions and in our case they give quite different answers.

We use daily data relating to three different markets which either trade or have their principal trading activities in different time zones. This data structure implies an overlapping structure for the price change data which requires modification of the standard VAR modelling approach. We have considered two different resolutions of this problem. The first involves deletion of leading lagged variables which overlap the disturbances on other questions. This approach maintains the price discovery interpretation in terms of markets but on our data it yields poorly determined error correction coefficients and hence possibly unreliable discovery shares. The alternative approach is to reformulate the VAR as a recursive structural VAR (SVAR). In this case, the relevant coefficients are better determined but the resulting discovery shares relate to the time slots defined by the determination of market closing prices rather than to the markets themselves.

Our substantive results differ depending on whether we rely on the PT or the IS discovery measure and depending on the VAR methodology. The standard VAR approach allows direct inference about the shares of the exchanges in price discovery. The estimates imply a predominant role for Comex and the SHFE in that order. The LME is seen as unimportant. However, these estimates rest on poorly determined error correction coefficients. We conclude that, in the absence of synchronous data, it is difficult to arrive at a judgement about the relative role of the different markets, as distinct form time slots, in price discovery. Our focus therefore shifts to time zones.

Using the recursive SVAR approach, the PT measure allocates discovery in a roughly equal manner to the New York and Shanghai trading periods with the London trading period contributing relatively little. Interpreting the PT measure as defining the best benchmark, these estimates again fail to support the practice of regarding LME prices as the best measure of the underlying fundamental copper price. The IS measure, which relates to information arrival, again indicate predominant roles for the New York and Shanghai trading slots with the London trading slot contributing even less than on the PT measure. There is little evidence that these rankings have varied over the twelve year period we have analyzed. It is possible that the relative unimportance of London daytime trading for price discovery in copper is less a reflection of the LME itself as of its location in a continent which has become relatively unimportant in terms of the world copper industry.

The LME has played a dominant role in non-ferrous metals futures since the nineteen fifties. Only in copper has the U.S., through Comex, been able to match a serious challenge. However, our estimates suggest that this reputation may reflect the LME's past success rather than its current importance. Europe's role in the world copper economy has declined over the past three decades. The LME responded to this challenge, first by moving out of its British base and opening warehouses in continental Europe and then by opening warehouses in Asia and North America. However, current regulations do not allow direct access to mainland China. This is the context of the 2012 purchase of the LME by the HKEx (Hong Kong Exchanges and Clearing). If HKEx obtains permission to open mainland Chinese warehouses, this will allow HKEx-LME to compete directly with the SHFE in the Chinese market.

This strategy may nevertheless be problematic. So long as the Chinese market remains only partially integrated with the world market, the LME will need to choose between a contract which competes directly with the SHFE and prices copper for mainland Chinese delivery and a contract which prices copper on the world market. Our estimates show that North American trading is of comparable or greater importance to that in China in terms of price discovery and is much more important than European trading. A decision to compete directly with the SHFE may end up in delivering the international price to Comex which has not currently declared Chinese ambitions.

References

Baillie, R.T., Booth, G.G., Tse, Y. and Zabotina, T, (2002). A theory of intraday patterns: volume and price variability. Review of Financial Studies, 1, 3-40.

Banks, J. 2011. China's rising taste for copper as collateral. The Banker, Special Report, Financial Market Series: The Commodity Supercycle.

Engle, R.F., and Granger. C.W.J. (1987), Cointegration and error correction: Representation, estimation and testing. Econometrica, 55, 251-76.

Figuerola-Ferretti, I., and Gonzalo, J., 2010. Modelling and Measuring Price Discovery in Commodity Markets. Journal of Econometrics, 158, 95-107.

Gonzalo, J., and Granger, C.W.J. (1995). Estimation of common long-memory components in cointegrated systems. Journal of Business and Economic Statistics, 13, 27-36.

Hasbrouck, J. 1995. One security, many markets: Determining the contributions to price discovery. Journal of Finance, 50, 1175-1199.

Hua, R., Lu, B. and Chen, B. (2010). "Price discovery process in copper markets: Is Shanghai futures market relevant?", Review of Futures Markets, 18.

Hupperets, E. C., and Menkveld, A.J. (2002). Intraday analysis of market integration: Dutch blue chips traded in Amsterdam and New York. Journal of Financial Markets, 5, 57-82.

Lieberman, O., Ben-Zion, U., and Hauser, S. (1999). A characterization of the price behavior of international dual stocks: an error correction approach. Journal of International Money and Finance, 18, 289-304.

Lien D. and Shrestha, K. (2009). A new information share measure. Journal of Futures Markets, 29, 377-395.

Lien D. and Shrestha, K. (2014). Price discovery in interrelated markets. Journal of Futures Markets, 34, 203-219.

Morrison, D.F. (1976). Multivariate Statistical Methods (2nd edition). New York: McGraw-Hill.

Pascual, R., Pascual-Fuster, B., and Climent, F. (2006). Cross-listing, price discovery and the informativeness of the trading process. Journal of Financial Markets, 9, 144-161.